Lecture 2 Exercises

Due Wed 9/12 at the beginning of class

1. Help the class collect a data set about books.

   a. First, find any two *hardcover* books of your choosing, check the number of pages x, and use the following function to predict the weight y in lbs.:

      $y = h_w(x) = .0024x + 1.5841$

   b. Then, look up each book on Amazon, and scroll down on the product information page for the book to find the actual shipping weight. Show how to use the actual shipping weight and the predicted weight to compute the value of the cost function $J(w_0, w_1) = J(1.5841, .0024)$ for these two books.

   c. Now, select a total of five books, and record the following data about them in the Google spreadsheet "CS 341 Book Data" (check your email for the link):

      i. Thickness (in.)
      ii. Width (in.)
      iii. Height (in.)
      iv. # Pages
      v. Shipping weight (lbs.)
      vi. Hardcover or paperback

      You can get all of this information from the Amazon product information page for the book. The dimensions are not always listed in the same order, so use the convention that *Thickness* < *Width* < *Height.*

2. In this exercise, you are going to solve a linear regression problem using two different approaches. Here is a small training set:

x	y
1	3
-1	-2
2	4

   Assume that $w_0 = 0$, so that $h_w(x) = w_1 x$, and

   $$J(w_1) = \sum_{i=1}^{N} (y_i - w_1 x_i)^2$$

   a. First, substitute the values of x_i and y_i from the training data into $J(w_1)$, and write the derivative. Set it equal to zero and solve for w_1.

   b. Second, use the general formula we derived at the end of class to find w_1.